人机对话方法、装置、设备及存储介质

本公开要求于 2021 年 09 月 30 日提交中国专利局、申请号为 202111165469.7、申请名称为“人机对话方法、装置、设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及信息技术领域，尤其涉及一种人机对话方法、装置、设备及存储介质。

背景技术

随着科技的不断发展，人机会话已经越来越普遍。

但是，本申请的发明人发现，表格知识是当前世界上存在最广泛的结构化知识，比如日常办公中的表格、网页中的各种表格、存在于多数企业和组织中的结构化查询语言（Structured Query Language，SQL）类的数据库等。因此，如何能让用户以人机对话的方式获取到表格知识中的内容就成为现有技术中亟需解决的问题。

发明内容

为了解决上述技术问题或者至少部分地解决上述技术问题，本公开提供了一种人机对话方法、装置、设备及存储介质，实现了使用户以人机对话的方式获取到表格知识的目的。

第一方面，本公开实施例提供一种人机对话方法，包括：

获取本轮的用户语句；

从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格；

对本轮的用户语句进行解析，得到第一查询语句；

根据第一查询语句，从一个或多个目标表格中查询获得目标数据；

根据目标数据生成本轮的系统回复。

第二方面，本公开实施例提供一种人机对话装置，包括：

获取模块，用于获取本轮的用户语句；

确定模块，用于从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格；

解析模块，用于对本轮的用户语句进行解析，得到第一查询语句；

查询模块，用于根据第一查询语句，从一个或多个目标表格中查询获得目标数据；

生成模块，用于根据目标数据生成本轮的系统回复。

第三方面，本公开实施例提供一种电子设备，包括：

存储器；

处理器；以及

计算机程序；

其中，所述计算机程序存储在所述存储器中，并被配置为由所述处理器执行以实现如第一方面所述的方法。

第四方面，本公开实施例提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行以实现第一方面所述的方法。

第五方面，本公开实施例提供一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现第一方面所述的方法。

本公开实施例提供的人机对话方法、装置、设备及存储介质，在获取到本轮对话的用户语句时，从一个或多个候选表格中确定与所述用户语句匹配的一个或多个目标表格，进而根据用户语句对应的第一查询语句从一个或多个目标表格中查询以获得目标数据，最后根据目标数据生成本轮对话中针对用户语句的系统回复；实现了基于表格知识为用户提供系统回复的目的。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

为了更清楚地说明本公开实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，对于本领域普通技术人员而言，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本公开实施例提供的人机对话方法流程图；

图 2 为本公开实施例提供的基于表格知识的人机对话交互系统的结构示意图；

图 3 为本公开实施例提供的自动生成训练数据的流程示意图；

图 4 为本公开实施例提供的基于表格知识的人机对话交互系统核心引擎的工作流程；

图 5 为本公开另一实施例提供的人机对话方法流程图；

图 6 为本公开另一实施例提供的人机对话方法流程图；

图 7 为本公开另一实施例提供的人机对话的交互界面示意图；

图 8 为本公开实施例提供的人机对话装置的结构示意图；

图 9 为本公开实施例提供的电子设备实施例的结构示意图。

具体实施方式

为了能够更清楚地理解本公开的上述目的、特征和优点，下面将对本公开的方案进行进一步描述。需要说明的是，在不冲突的情况下，本公开的实施例及实施例中的特征可以相互组合。

在下面的描述中阐述了很多具体细节以便于充分理解本公开，但本公开还可以采用其他不同于在此描述的方式来实施；显然，说明书中的实施例只是本公开的一部分实施例，而不是全部的实施例。

通常，人机对话基于知识实现，该知识包括：流程类的知识、图谱类的知识、以及文档。然而，随着表格知识的不断发展，目前，二维表格知识是当前世界上存在较广泛的结构化知识，比如日常办公中的表格、网页中的各种表格、存在于多数企业和组织中的 SQL 类的数据库等。而且，表格知识结构清晰、易于维护、时效性强。因此，如何让人以自然语言对话的方式和这些表格知识进行交互，是工业界和学术界共同关注的热点技术方向，也是人工智能研究的前沿方向之一。针对该问题，本公开实施例提供了一种基于表格知识的人机对话方法，下面结合具体的实施例对该方法进行介绍。

图 1 为本公开实施例提供的人机对话方法流程图。本实施例可适用于客户端中进行人机对话的情况，该方法可以由人机对话装置执行，该装置可以采用软件和/或硬件的方式实现，该装置可配置于电子设备中，例如终端，具体包括手机、电脑或平板电脑等。或者，本实施例可适用于服务端中进行人机对话的情况，该方法可以由人机对话装置执行，该装置可以采用软件和/或硬件的方式实现，该装置可配置于电子设备中，例如服务器。

另外，本实施例所述的人机对话方法可以适用于多种场景，例如，用户和终端进行人机对话，该终端可以是机器人、手机、电脑等用户设备。在用户和机器人对话的场景中，用户和机器人可以在同一地点进行面对面对话，或者用户和机器人在不同地点进行远程对话。用户和终端的对话方式不限于语音对话，还可以通过文字、手势等方式进行对话。终端接收到用户的提问（即用户语句）后，需要根据该提问给出相应的回复，该回复在本公开实施例中称为系统回复。

参考如图 2 所示的一种基于表格知识的人机对话交互系统的结构示意图，该人机对话交互系统包括四大部分，分别是数据和知识部分 210、预训练模型部分 220、核心引擎部分 230 和应用层部分 240。

其中，数据和知识部分 210 包括数据合成器 211 和表知识库 table212。表知识库 table212 中存储的是某种类型的表格数据，例如关于车辆的表格数据、关于某业务办理流程的表格数据等。表格数据的具体类型可根据应用场景确定。例如应用场景为汽车平台的智能客服系统，表格数据的类型可以是车辆相关，表格数据可以是关于车辆的表格数据。数据合成器 211 的功能是自动合成<Text, SQL> 数据对，即自动合成自然语言语句 Text 对应的查询语句 SQL。生成的大量数据对用于作为核心引擎部分 230 中语言理解装置 231 的训练数据集，后续进行进一步说明。

预训练模型部分 220，在自然语言处理领域，预训练语言模型的出现是一个重大突破，给下游任务带来了普遍的、显著的效果提升。在本公开的技术方案中，预训练模型采用面向表格知识的预训练表格模型，该预训练表格模型具体包括预训练表格理解模型、预训练表格生成模型等。

核心引擎部分 230 包括语言理解装置 231、对话管理装置 232 和语言生成装置 233。

其中，语言理解装置 231 的主要功能是提供一种语言理解的方式 Text-to-SQL，将用户使用自然语言输入的问句 Text 解析为机器可以识别的查询语句 SQL，将该查询语句 SQL 记为第一查询语句。可选的，语言理解装置 231 将用户的问句 Text 解析为第一查询语句的过程如下：语言理解装置 231 在接收到用户的当前问句后，从候选表格中挑选出满足当前问句的一个或多个表格，将挑选出的满足当前问句的一个或多个表格记为目标表格；进一步将目标表格和当前问句一起作为 Text-to-SQL 模型的输入，Text-to-SQL 模型输出的 SQL 语句即为当前问句对应的第一查询语句。语言理解装置 231 在得到第一查询语句后，进一步将第一查询语句提供给对话管理装置 232，也可以是对话管理装置 232 主动从语言理解装置 231 请求第一查询语句。

对话管理装置 232 的主要功能是提供一种语言合成的方式 SQL-to-Text，将 SQL 语句转换成自然语言语句 Text、管理多轮对话中的对话状态，并根据对话状态给出 SQL 语句，将该 SQL 语句记为第二查询语句，进一步根据第二查询语句从上述目标表格中查询相关表格数据，将查询到的相关表格数据记为查询结果。可选的，将查询结果提供给语言生成装置 233，或者语言生成装置 233 主动从对话管理装置 232 请求查询结果。需要特别说明的是，第二查询语句与通过语言理解装置 231 得到的第一查询语句不同。具体的，在对话管理装置 232 获取到第一查询语句后，根据下述三个因素中的

至少一个对第一查询语句进行修正，获得当前问句对应的第二查询语句；所述三个因素具体是：用户的当前问句之前的上一轮对话中的用户历史问句对应的修正后的 SQL 语句（即上一轮对话中的第二查询语句）、当前问句对应的会话动作和上一轮对话所对应的系统状态。其中，“用户的当前问句之前的上一轮对话中的用户历史问句对应的修正后的 SQL 语句”存储在图 4 所示的“上下文”的数据库中。“当前问句对应的会话动作”是对话管理装置 232 根据上一轮对话中的第二查询语句和当前轮对话中的第一查询语句得到的。“上一轮对话所对应的系统状态”是指：在上一轮对话中，系统是直接回复了话术，还是反问了用户。“上一轮对话所对应的系统状态”也存储在图 4 所示的“上下文”的数据库中。用户的一次提问和系统针对该提问给出的回复作为一轮对话。“会话动作”包括：增加查询条件、删除查询条件和修改查询条件。具体的，增加查询条件例如是：用户输入第一个问句：“收益率在 3.5%以上的理财产品有哪些？”，系统针对第一个问句进行回复后，用户接着输入第二个问句“那保本的呢？”，此场景下用户输入的第二个问句其实是在第一个问句的基础上增加了查询条件，即第二个问句对应的完整信息是“收益率在 3.5%以上且保本的理财产品有哪些”，在查询条件“收益率在 3.5%以上”的基础上增加了查询条件“保本”。删除查询条件例如是：用户输入第一个问句“能保本的且收益率在 3.5%以上的理财产品有哪些？”，系统针对第一个问句进行回复后，用户又输入第二个问句“那收益率在 3.5%以上的呢？”，此场景下用户的第二个问句其实是在第一个问句的基础上删除了查询条件，即第二个问句对应的完整信息是“收益率在 3.5%以上的理财产品有哪些”。修改查询条件例如是：用户输入第一个问句“某品牌的轿车有没有天窗？”，系统针对该第一个问句进行回复后，用户又输入第二个问句“某品牌的越野车性能如何？”，此场景下用户的第二个问句对应的查询条件与第一个问句对应的查询条件不同，为修改查询条件的场景。进一步举例说明对第一查询语句进行修正，获得当前问句对应的第二查询语句的过程：

用户问：太阳牌的轿车有没有天窗？（该问句为历史问句）

系统回复：太阳牌的轿车有天窗。（用户的一次提问和系统给出的针对性回复组成一轮对话）

用户又问：那百公里加速呢？（该问句为当前问句）

其中，问句“太阳牌的轿车有没有天窗？”对应的 SQL 语句例如为：“SELECT 天窗 WHERE 太阳牌轿车”。问句“那百公里加速呢？”对应的 SQL 语句（即第一查询语句）为：“SELECT 百公里加速”。对第一查询语句修正后得到的第二查询语句为“SELECT

百公里加速 WHERE 太阳牌轿车”。进一步，对话管理装置 232 根据第二查询语句到目标表格中进行数据查询，查询到的数据记为查询结果，并将查询结果提供给语言生成装置 233。

语言生成装置 233 的主要功能是提供一种语言生成的方式 Data-to-Text，将表格数据 Data 转换成自然语言语句 Text，具体是根据用户的当前问句、第二查询语句和查询结果，生成回复话术，具体是使用自然语言针对用户的当前问句给出针对性的回复。例如，用户的当前问句是：某品牌的轿车有没有天窗，查询结果是 Y，Y 代表有天窗，则生成的回复话术为：某品牌的轿车有天窗。

应用层部分 240，基于表格知识的人机对话交互系统有很多应用，典型的比如面向对话机器人的多轮问答系统和面向 SQL 数据库的自然交互接口。其中，面向对话机器人的多轮问答系统的应用场景例如是人和机器人对话，人和机器人在同一地点进行面对面对话，或在不同地点进行远程对话。面向 SQL 数据库的自然交互接口的应用场景例如是用户在浏览器或用户界面中输入问句，终端将该问句发送给远程服务器，远程服务器内集成有 SQL 数据库，或者远程服务器与 SQL 数据库通信连接；进一步，远程服务器根据用户问句从 SQL 数据库所存储的表格中查询到数据，将该数据发送给终端，终端根据该数据生成自然语言，并向用户给出回复。

为了充分公开本实施例的技术方案，针对图 2 所示的基于表格知识的人机对话交互系统中的关键部分进行进一步说明。

具体的，基于表格知识的人机对话交互系统依托表格知识完成系统与用户的交互，其核心功能是如何将用户使用自然语言输入的问句（即用户的提问 Text）转化成 SQL 表达式（即机器可以识别的 SQL 查询语句），该核心功能的实现依赖 Text-to-SQL 模型，即核心引擎部分 230 中的语言理解装置 231。

Text-to-SQL 模型的训练需要大量的<Text, SQL> 数据对作为训练数据，但通过人工标注获得训练数据的难度较大、成本较高，针对该问题，本公开提出一种自动生成训练数据的数据合成器 211，在给定表格知识时，通过数据合成器 211 自动生成训练数据，其自动生成训练数据的流程如图 3 所示。

具体的，给定表格知识，表格知识指以表格形式存储的数据；首先对表格知识进行预处理，包括归一化、特殊字符的处理等；通过预处理得到符合规范的表格知识。然后通过 SQL 采样器，生成与给定的候选表格相关的大量 SQL 语句，具体生成过程如下：SQL 采样器中可以预先设置一个或多个 SQL 模板，例如，其中一个 SQL 模板为：SELECT A，WHERE B；进一步，将该 SQL 模板中的 A 和 B 分别替换为给定的候选表格中

的字段；由于候选表格中存在多个字段，例如，列名、列值等，因此，可以将 A 替换为该候选表格中的某个字段，将 B 替换为该候选表格中的另一个字段，从而得到多个实例化的 SQL 语句。进一步，还可以将 A 和 B 分别替换成的字段进行更新，从而得到又一组实例化的 SQL 语句。依次类推，可得到给定候选表格对应的很多实例化的 SQL 语句。概括性的，根据 SQL 模板和某个候选表格得到 SQL 语句，基于不同的候选表格得到不同的 SQL 语句。最后将每个 SQL 语句转换成对应的自然语言，例如，一个 SQL 语句对应一个自然语言。为了提高话术的泛化性和生成话术的流畅性，可以将每个 SQL 语句转换成的自然语言中的一些词替换为该词的同义词或同义短语等，从而使得一个 SQL 语句可以对应多个自然语言，从而增加了训练数据的多样性和数量，降低了训练数据的获取难度。例如，将某个 SQL 语句转换成对应的自然语言是“年化收益率为 3.5%”，由于“年化收益率”和“收益率”是同义词，因此，该 SQL 语句对应的自然语言可以有两个，一个是“年化收益率为 3.5%”，另一个是“收益率为 3.5%”。进一步，根据该 SQL 语句和其对应的两个自然语言，可以得到两组数据对<Text, SQL>，其中一组数据对<Text, SQL>中的 Text 是“年化收益率为 3.5%”；另一组数据对<Text, SQL>中的 Text 是“收益率为 3.5%”，而这两组数据对<Text, SQL>中的 SQL 为同一 SQL 语句。由于与给定的候选表格相关的 SQL 语句有很多，其中一个 SQL 语句可以对应有一个或多个自然语言语句 Text，因此，可以得到大量的数据对<Text,SQL>。

当前自然语言处理领域的任务，基本范式是 pretrain（预训练）+finetune（微调），即以预训练模型（典型的比如 BERT 模型）为基础模型，进一步采用具体任务的标注数据即根据某个或某几个给定的表格知识得到的数据对<Text, SQL>，在基础模型上做 finetune（微调）即对基础模型的参数进行微调。在本实施例中，该基础模型选用预训练表格模型，而不是预训练语言模型，这是因为预训练语言模型有两个缺点：（1）预训练语言模型没有将文本数据和表格知识结合；（2）预训练语言模型主要学习语言知识（例如语句是否通顺、语法是否正确），缺乏对结构化知识的建模。所以本案采用专门面向表格知识的预训练表格模型作为基础模型。

基于表格知识的人机对话交互系统核心引擎的工作流程如图 4 所示。首先用户使用自然语言发出提问，比如输入问句 Text “收益率 3.5%以上的理财产品有哪些？”，语言理解装置 231 接收到用户语句 Text 后，先从所有候选表格（候选表格可以是一个或多个，存储在 Table 库 212 中）中选出满足用户问句 Text 的表格，将满足用户问句 Text 的表格标记为目标表格。例如，Table 库 212 中默认有 3 张表格，分别是基金表格、证券表格和外汇表格，每个表格中的内容不同，最能满足用户问句 Text 的表格可

以是该 3 张表格中的一个或多个。从候选表格中选出目标表格的过程有如下几种可选方法：一种是选取用户问句 Text 中的关键词，将该关键词与每个表格的描述信息进行匹配，选出匹配度达到阈值的一个或多个表格作为目标表格。另一种是通过一个预先训练好的模型来选，例如，该模型的输入是用户问句 Text，该模型的输出是用户问句 Text 与每个候选表格的匹配度，进一步将匹配度达到阈值的一个或多个候选表格作为目标表格。然后将目标表格、用户问句 Text 作为 Text-to-SQL 模型的输入，Text-to-SQL 模型输出用户问句 text 对应的 SQL 语句（即第一查询语句），该 SQL 语句例如为“SELECT 产品名称 WHERE 收益率 == 3.5%”。然后对话管理装置 232 获取到语言理解装置 231 的处理结果（即语言理解装置 231 得到的 SQL 语句，也就是图 4 中（2）所示的 SQL），可以是对话管理装置 232 主动从语言理解装置 231 获取处理结果，也可以是对话管理装置 232 被动接收语言理解装置 231 发送的处理结果。进一步，对话管理装置 232 对该处理结果继续进行处理，换言之，对话管理装置 232 对第一查询语句进行修正获得第二查询语句（即图 4 中（3）所示的 SQL），修正过程可参见前述说明。图 4 中（3）所示的 SQL 与（2）所示的 SQL 有所不同：（2）所示的 SQL 是语言理解装置 231 根据用户当前的问句 Text 例如“收益率 3.5%以上的理财产品有哪些”得到的 SQL 语句，但（3）所示的 SQL 是对（2）所示的 SQL 进行修正后得到的 SQL 语句。（3）所示的 SQL 是最终的、更完整的 SQL 语句。对话管理装置 232 将生成的修改后的 SQL，即（3）所示的 SQL 发送到表格知识库（即语言理解装置 231 从候选表格中选出的一个或多个目标表格）进行知识查询，得到查询结果（从目标表格中查询到的数据），对话管理装置 232 进一步将修正后的 SQL 即图 4 中（3）所示的 SQL 以及该查询结果（该查询结果对应图 4 中的（5）所示的对话管理结果）发送给语言生成装置 233；最后语言生成装置 233 基于修改后的 SQL、该查询结果以及用户当前的问句 Text（语言生成装置 233 从图 4 中上下文的数据库中提取到用户当前的问句 Text）生成回复话术“收益率大于 3.5%的理财产品有 a、b 等 10 款产品”，并回复给用户，至此在线流程完成一轮完整的交互。可以理解的是，图 4 中的语言理解装置、对话管理装置、存储上下文的存储空间、语言生成装置、存储表格知识的存储空间等可以部署在同一个服务器，也可以部署在不同的服务器上，如果部署在多个不同的服务器上，则多个不同的服务器可以构成服务器集群。图 4 中对话管理装置向上下文数据库中写入上下文的含义是：将用户当前的问句 Text、用户当前的问句 Text 对应的修正后的 SQL 语句以及本轮对话对应的系统状态写入上下文数据库中，为下一轮对话提供参考数据。

概括性的，在用户基于终端进行人机对话的场景中，终端根据用户的问句确定系

统回复的过程中，需要用到图 2 所示的基于表格知识的人机对话交互系统中的核心引擎部分 230 中的 3 个装置，即语言理解装置 231、对话管理装置 232 和语言生成装置 233。核心引擎部分 230 可以部署在终端本地，或者也可以部署在与终端能够通信的远程服务器上，或者核心引擎部分 230 中的部分装置部署在终端本地、部分装置部署在远程服务器上。对于核心引擎部分 230 中的 3 个装置或部分装置部署在远程服务器上的这种情况，该 3 个装置或部分装置可以部署在同一台远程服务器上，也可以分别部署在不同的远程服务器上。另外，针对核心引擎部分 230 中的同一个装置，该装置还可以部署在多台远程服务器上，由多台远程服务器实现该装置对应的功能。此外，如图 2 所示的核心引擎部分 230 中的语言理解装置 231 具体是 Text-to-SQL 模型。如图 2 所示的数据和知识部分 210 和预训练模型部分 220 的作用是对 Text-to-SQL 模型进行训练。数据和知识部分 210 的作用是提供训练数据即训练样本数据。预训练模型部分 220 的作用是提供预训练表格模型，该预训练表格模型可以作为基础模型，基础模型是预先通过广泛的、大量的、多种不同类型的表格数据训练得到的。而数据和知识部分 210 提供的训练样本数据是根据某种类型的表格数据生成的，某种类型的表格数据例如为关于车辆信息的表格数据。通过数据和知识部分 210 提供的训练样本数据对该基础模型进行训练，得到 Text-to-SQL 模型。可以理解的是，在没有基础模型的情况下，通过数据和知识部分 210 提供的训练样本数据也可以直接训练出 Text-to-SQL 模型，但是，在这种情况下，相当于 Text-to-SQL 模型没有对应的基础模型，Text-to-SQL 模型的参数需要从 0 开始优化，从而导致训练周期较长。因此，数据和知识部分 210 和预训练模型部分 220 可以部署在远程服务器上，例如，数据和知识部分 210 和预训练模型部分 220 可以部署在同一个或不同的远程服务器上。当 Text-to-SQL 模型在远程服务器上被训练完成后，Text-to-SQL 模型可以被移植到终端，由终端实现语言理解的相关功能。或者，Text-to-SQL 模型可以在某一个远程服务器上被训练完成后，移植到其他远程服务器上，由其他远程服务器实现语言理解的相关功能。

进一步的，下面以终端为例介绍该人机对话方法，该终端包括但不限于智能手机、掌上电脑、平板电脑、带显示屏的可穿戴设备、台式机、笔记本电脑、一体机、智能家居设备等。如图 1 所示，该方法具体步骤如下：

S101、获取本轮的用户语句。

其中，用户语句为用户主动输入，通常是用户的提问，例如用户语句是“收益率在 3.5%以上的理财产品有哪些？”。用户输入用户语句的方式可以是语音输入、文字

输入或者手势输入等。

通常将用户输入的用户语句以及系统针对该用户语句给出的回复确定为一轮人机对话。本轮的用户语句指当前的用户问句，换言之，本轮的用户语句指距离当前时刻最近的一次用户的提问。

具体的，用户通过终端输入用户语句，基于表格知识的人机对话交互系统中的语言理解装置 231 获取到用户语句。

S102、从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格。

具体的，语言理解装置 231 获取到用户语句后，先从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格。其中，一个或多个候选表格存储在图 2 中的 Table 库 212 中。例如，Table 库 212 中默认有 3 张表格，分别是基金表格、证券表格和外汇表格，每个表格中的内容不同，最能满足用户语句的表格可以是该 3 张表格中的一个或多个。

从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格包括：

确定用户语句中的关键词，将关键词与每个候选表格的描述信息进行匹配，将匹配度达到阈值的一个或多个表格作为目标表格。或者，通过预先训练好的模型来选，具体的，将用户语句输入至预先训练好的模型，获得用户语句与每个候选表格的匹配度，将匹配度达到阈值的一个或多个候选表格作为目标表格。

S103、对本轮的用户语句进行解析，得到第一查询语句。

可选的，对本轮的用户语句进行解析，得到第一查询语句，包括：将一个或多个目标表格、以及本轮的用户语句作为预设模型的输入，通过预设模型得到第一查询语句。预设模型具体可以是上述的 Text-to-SQL 模型。

例如用户语句是“收益率在 3.5%以上的理财产品有哪些？”，对应的第一查询语句是 “SELECT 产品名称 WHERE 收益率 == 3.5%”。

进一步的，也可以通过目前被广泛使用的将自然语言转换为 SQL 语言的方式对本轮的用户语句进行解析，得到第一查询语句。

S104、根据第一查询语句，从一个或多个目标表格中查询获得目标数据。

可选的，通过语言理解装置 231 获得第一查询语句后，语言理解装置 231 将第一查询语句发送给对话管理装置 232，或者对话管理装置 232 主动从语言理解装置 231 获取第一查询语句。进一步，对话管理装置 232 对第一查询语句进行修正获得第二查询语句，并根据第二查询语句从一个或多个目标表格中查询获得目标数据。进一步将目标数据以及第二查询语句提供给语言生成装置 233。

进一步的，对话管理装置 232 在获取到第一查询语句时，也可直接基于第一查询语句从一个或多个目标表格中查询获得目标数据。进一步将目标数据以及第二查询语句提供给语言生成装置 233。可以理解的是，参考的信息维度越多，则从目标表格中查询获得的目标数据越准确，越符合用户的需求。

S105、根据目标数据生成本轮的系统回复。

在一些实施方式中，语言生成装置 233 基于目标数据以及本轮的用户语句（语言生成装置 233 从图 4 中上下文的数据库中提取到用户语句）生成系统回复，例如是“收益率大于 3.5%的理财产品有 a、b 等 10 款产品”，并回复给用户，向用户回复的方式可以是通过语音播放给用户，或者通过终端显示屏进行显示等。

在另一些实施方式中，语言生成装置 233 基于目标数据、本轮的用户语句以及第二查询语句（语言生成装置 233 从图 4 中上下文的数据库中提取到用户语句）生成系统回复。通过参考更多的相关信息可使得生成的系统回复更通顺，可读性更好。

本公开实施例在获取到本轮对话的用户语句时，从一个或多个候选表格中确定与所述用户语句匹配的一个或多个目标表格，进而根据用户语句对应的第一查询语句从一个或多个目标表格中查询以获得目标数据，最后根据目标数据生成本轮对话中针对用户语句的系统回复；实现了基于表格知识为用户提供系统回复的目的。

图 5 为本公开另一实施例提供的人机对话方法流程图。在本实施例中，针对上述步骤 S104 给出了一种可选实施方式。如图 5 所示，该方法具体步骤如下：

S501、获取本轮的用户语句。

S502、从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格。

S503、对本轮的用户语句进行解析，得到第一查询语句。

S504、根据历史对话数据中用户历史语句对应的历史查询语句、本轮的用户语句对应的会话动作、历史对话数据对应的系统状态中的至少一个，对第一查询语句进行处理，得到第二查询语句。

其中，会话动作是根据历史查询语句和第一查询语句得到的；历史对话数据对应的系统状态用于表征历史对话数据中的系统历史回复是否为反问语句。

历史对话数据指用户与系统的历史对话记录，历史对话数据包括一轮或多轮人机对话，每轮人机对话中的用户语句为用户历史语句，每个用户历史语句均对应有历史查询语句，即 SQL 语句。

具体的，在对话管理装置 232 获取到第一查询语句后，根据下述三个因素中的至少一个对第一查询语句进行修正，获得本轮用户语句对应的第二查询语句。所述三个

因素具体是：历史对话数据中用户历史语句对应的历史查询语句、本轮的用户语句对应的会话动作和历史对话数据对应的系统状态。其中，历史对话数据中用户历史语句对应的历史查询语句指：本轮对话之前的上一轮对话中的用户语句对应的修正后的SQL语句（即上一轮对话中的第二查询语句），该历史查询语句存储在图4所示的“上下文”的数据库中；本轮的用户语句对应的会话动作是对话管理装置232根据上一轮对话中的第二查询语句和本轮对话中的第一查询语句得到的，即会话动作是根据历史查询语句和第一查询语句得到的。“历史对话数据对应的系统状态”是指：在上一轮对话中，系统是直接回复了话术，还是反问了用户。“历史对话数据对应的系统状态”也存储在图4所示的“上下文”的数据库中。用户的一次提问和系统针对该提问给出的回复作为一轮对话。“会话动作”包括：增加查询条件、删除查询条件和修改查询条件。具体的，增加查询条件例如是：用户输入第一个语句：“收益率在3.5%以上的理财产品有哪些？”，系统针对第一个语句进行回复后，用户接着输入第二个语句“那保本的呢？”，此场景下用户输入的第二个语句其实是在第一个语句的基础上增加了查询条件，即第二个语句对应的完整信息是“收益率在3.5%以上且保本的理财产品有哪些”，在查询条件“收益率在3.5%以上”的基础上增加了查询条件“保本”。删除查询条件例如是：用户输入第一个语句“能保本的且收益率在3.5%以上的理财产品有哪些？”，系统针对第一个语句进行回复后，用户又输入第二个语句“那收益率在3.5%以上的呢？”，此场景下用户的第二个语句其实是在第一个语句的基础上删除了查询条件，即第二个语句对应的完整信息是“收益率在3.5%以上的理财产品有哪些”。修改查询条件例如是：用户输入第一个语句“某品牌的轿车有没有天窗？”，系统针对该第一个语句进行回复后，用户又输入第二个语句“某品牌的越野车性能如何？”，此场景下用户的第二个语句对应的查询条件与第一个语句对应的查询条件不同，为修改查询条件的场景。进一步举例说明对第一查询语句进行修正，获得当前问句对应的第二查询语句的过程：

用户语句：太阳牌的轿车有没有天窗？（该语句为用户历史语句）

系统回复：太阳牌的轿车有天窗。（用户的一次提问和系统给出的针对性回复组成一轮对话）

用户又问：那百公里加速呢？（该语句为本轮的用户语句）

其中，历史用户语句“太阳牌的轿车有没有天窗？”对应的历史查询语句例如为：“SELECT 天窗 WHERE 太阳牌轿车”。本轮的用户语句“那百公里加速呢？”对应的SQL语句（即第一查询语句）为：“SELECT 百公里加速”。对第一查询语句修正后得

到的本轮对话中的第二查询语句为“SELECT 百公里加速 WHERE 太阳牌轿车”。通过对第一查询语句进行修正获得的第二查询语句与用户的查询意图更加匹配，即更能体现用户的查询意图，进而使得基于第二查询语句获得的目标数据更准确，使得最后的系统回复更具针对性，实现所答即所问的技术效果。

S505、根据第二查询语句，从一个或多个目标表格中查询获得目标数据。

进一步，对话管理装置 232 根据第二查询语句到目标表格中进行数据查询，查询到的数据记为目标数据，并将目标数据提供给语言生成装置 233。

S506、根据目标数据生成本轮的系统回复。

语言生成装置 233 根据目标数据和本轮的用户语句，生成本轮的系统回复。例如，本轮的用户语句是：太阳牌的轿车有没有天窗，目标数据是 Y，Y 代表有天窗，则语言生成装置 233 根据目标数据和本轮的用户语句，生成本轮的系统回复为：太阳牌的轿车有天窗。

具体的，S501-S503 和 S101-S103 的实现方式和具体原理一致，此处不再赘述。

本实施例针对上述步骤 S104 给出了一种可选实施方式，具体是在根据第一查询语句，从一个或多个目标表格中查询获得目标数据的过程中，首先根据历史对话数据中用户历史语句对应的历史查询语句、本轮的用户语句对应的会话动作、历史对话数据对应的系统状态中的至少一个，对第一查询语句进行处理，得到第二查询语句，然后根据第二查询语句从一个或多个目标表格中查询获得目标数据，提高了目标数据的查询精度，使得基于表格知识为用户提供系统回复的效果更好。

图 6 为本公开另一实施例提供的一种人机对话方法的流程图。在本实施例中，增加了获取预设模型的步骤。如图 6 所示，该方法具体步骤如下：

S601、获取本轮的用户语句，从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格。

S602、生成一个或多个候选表格中每个候选表格相关的样本查询语句；将样本查询语句转换为第一自然语言语句；根据每个候选表格相关的样本查询语句、以及第一自然语言语句，得到训练数据。

具体的，可以通过 SQL 采样器，生成与每个候选表格相关的大量 SQL 语句，即样本查询语句。具体生成过程如下：SQL 采样器中可以预先设置一个或多个 SQL 模板，例如，其中一个 SQL 模板为：SELECT A，WHERE B；进一步，将该 SQL 模板中的 A 和 B 分别替换为候选表格中的字段；由于候选表格中存在多个字段，例如，列名、列值等，因此，可以将 A 替换为该候选表格中的某个字段，将 B 替换为该候选表格中的另一个

字段，从而得到多个实例化的 SQL 语句。进一步，还可以将 A 和 B 分别替换成的字段进行更新，从而得到又一组实例化的 SQL 语句。依次类推，可得到每个候选表格对应的很多实例化的 SQL 语句。概括性的，根据 SQL 模板和某个候选表格得到样本查询语句，基于不同的候选表格得到不同的样本查询语句。最后将每个样本查询语句转换成对应的第一自然语言语句，例如，一个样本查询语句对应一个自然语言语句。为了提高话术的泛化性和生成话术的流畅性，可以将每个样本查询语句转换成的第一自然语言语句中的一些词替换为该词的同义词或同义短语等，从而使得一个样本查询语句可以对应多个自然语言语句，从而增加了训练数据的多样性和数量，降低了训练数据的获取难度。例如，将某个样本查询语句转换成对应的自然语言语句是“年化收益率为 3.5%”，由于“年化收益率”和“收益率”是同义词，因此，该样本查询语句对应的自然语言语句可以有两个，一个是“年化收益率为 3.5%”，另一个是“收益率为 3.5%”。进一步，根据该样本查询语句和其对应的两个自然语言语句，可以得到两组数据对<Text, SQL>，其中一组数据对<Text, SQL>中的 Text 是“年化收益率为 3.5%”；另一组数据对<Text, SQL>中的 Text 是“收益率为 3.5%”，而这两组数据对<Text, SQL>中的 SQL 为同一样本查询语句。由于与给定的候选表格相关的样本查询语句有很多，其中一个样本查询语句可以对应有一个或多个自然语言语句 Text，因此，可以得到大量的数据对<Text,SQL>，将大量的数据对<Text,SQL>确定为训练数据。

概括性的，根据每个候选表格相关的样本查询语句、以及第一自然语言语句，得到训练数据，包括：将第一自然语言语句中的关键词替换为关键词对应的同义词或同义短语，得到样本查询语句对应的一个或多个第二自然语言语句；根据每个候选表格相关的样本查询语句、第一自然语言语句、以及一个或多个第二自然语言语句，得到训练数据。

S603、根据训练数据进行模型训练得到预设模型。

可选的，根据训练数据对预训练模型进行训练得到预设模型，预训练模型是预先根据表格信息训练得到的模型。

当前自然语言处理领域的任务，基本范式是 pretrain（预训练）+finetune（微调），即以预训练模型（典型的比如 BERT 模型）为基础模型，进一步采用具体任务的标注数据即根据某个或某几个给定的表格知识得到的数据对<Text, SQL>，在基础模型上做 finetune（微调）即对基础模型的参数进行微调。在本实施例中，该基础模型选用预训练表格模型，而不是预训练语言模型，这是因为预训练语言模型有两个缺点：（1）预训练语言模型没有将文本数据和表格知识结合；（2）预训练语言模型主要学

习语言知识（例如语句是否通顺、语法是否正确），缺乏对结构化知识的建模。所以本案采用专门面向表格知识的预训练表格模型作为基础模型。可以理解的是，在没有基础模型的情况下，通过训练数据也可以直接训练出预设模型，但是，在这种情况下，相当于预设模型没有对应的基础模型，预设模型的参数需要从 0 开始优化，从而导致训练周期较长。

S604、将一个或多个目标表格、以及本轮的用户语句作为预设模型的输入，通过预设模型得到第一查询语句。

S605、根据历史对话数据中用户历史语句对应的历史查询语句、本轮的用户语句对应的会话动作、历史对话数据对应的系统状态中的至少一个，对第一查询语句进行处理，得到第二查询语句，根据第二查询语句，从一个或多个目标表格中查询获得目标数据。

S606、根据目标数据生成本轮的系统回复。

本实施例增加了获取预设模型的步骤，具体是首先生成一个或多个候选表格中每个候选表格相关的样本查询语句，将样本查询语句转换为第一自然语言语句，根据每个候选表格相关的样本查询语句、以及第一自然语言语句，得到训练数据；然后根据训练数据对预训练模型进行训练得到预设模型，预训练模型是预先根据表格信息训练得到的模型。通过将每个样本查询语句转换成的第一自然语言语句中的一些词替换为该词的同义词或同义短语等，从而使得一个样本查询语句可以对应多个自然语言语句，从而增加了训练数据的多样性和数量，降低了训练数据的获取难度，实现了训练数据的自动生成；通过使用预训练模型作为基础模型，缩短了预设模型的训练周期，提升了预设模型的模型性能和效果；通过采用专门面向表格知识的预训练表格模型作为基础模型，实现了基于表格知识为用户提供系统回复的目的。

在上述各实施例的基础上，参考如图 7 所示的一种在具体应用场景中，人机对话的交互界面示意图，其中左边的消息是系统给出的系统回复，右边的消息是用户语句，在实际业务场景中，通过多轮对话可解决实际的用户问题。示例性的：

用户语句：我想了解一下户籍的办理途经。

系统回复：请选择具体业务：户口登记、恢复户口、注销户口。

用户语句：恢复户口。

系统回复：已为您找到具体业务是恢复户口的，请选择具体事项：军人退伍、出国出境公民回国申报、被判刑或者劳动教养的公民。

用户语句：军人退伍。

系统在对用户的第二个语句进行回复时，结合系统的第一次回复以及用户的第二个语句，即系统对用户进行第二次回复是基于对话历史确定的。

图 8 为本公开实施例提供的人机对话装置的结构示意图。本公开实施例提供的人机对话装置可以执行人机对话方法实施例提供的处理流程，如图 8 所示，该人机对话装置 80 包括：获取模块 810、确定模块 820、解析模块 830、查询模块 840 和生成模块 850。

其中，获取模块 810 用于获取本轮的用户语句；确定模块 820 用于从一个或多个候选表格中确定与用户语句匹配的一个或多个目标表格；解析模块 830 用于对本轮的用户语句进行解析，得到第一查询语句；查询模块 840 用于根据第一查询语句，从一个或多个目标表格中查询获得目标数据；生成模块 850 用于根据目标数据生成本轮的系统回复。

可选的，查询模块 840 包括处理单元，用于根据历史对话数据中用户历史语句对应的历史查询语句、本轮的用户语句对应的会话动作、历史对话数据对应的系统状态中的至少一个，对第一查询语句进行处理，得到第二查询语句；查询单元，用于根据第二查询语句，从一个或多个目标表格中查询获得目标数据；其中，会话动作是根据历史查询语句和第一查询语句得到的；历史对话数据对应的系统状态用于表征历史对话数据中的系统历史回复是否为反问语句。

可选的，生成模块 850 具体用于：根据目标数据和本轮的用户语句，生成本轮的系统回复。

可选的，解析模块 830 具体用于：将一个或多个目标表格、以及本轮的用户语句作为预设模型的输入，通过预设模型得到第一查询语句。

可选的，还包括：训练数据生成模块，用于在将一个或多个目标表格、以及本轮的用户语句作为预设模型的输入，通过预设模型得到第一查询语句之前，生成一个或多个候选表格中每个候选表格相关的样本查询语句；将样本查询语句转换为第一自然语言语句；根据每个候选表格相关的样本查询语句、以及第一自然语言语句，得到训练数据；训练模块，用于根据训练数据进行模型训练得到预设模型。

可选的，训练模块具体用于：根据训练数据对预训练模型进行训练得到预设模型，预训练模型是预先根据表格信息训练得到的模型。

可选的，训练数据生成模块包括：替换单元，用于将第一自然语言语句中的关键词替换为关键词对应的同义词或同义短语，得到样本查询语句对应的一个或多个第二自然语言语句；根据每个候选表格相关的样本查询语句、第一自然语言语句、以及一

个或多个第二自然语言语句，得到训练数据。

图 8 所示实施例的人机对话装置可用于执行上述方法实施例的技术方案，其实现原理和技术效果类似，此处不再赘述。

以上描述了人机对话装置的内部功能和结构，该装置可实现为一种电子设备。图 9 为本公开实施例提供的电子设备实施例的结构示意图。如图 9 所示，该电子设备包括存储器 151 和处理器 152。

存储器 151，用于存储程序。除上述程序之外，存储器 151 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 151 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器(SRAM)，电可擦除可编程只读存储器(EEPROM)，可擦除可编程只读存储器(EPROM)，可编程只读存储器(PROM)，只读存储器(ROM)，磁存储器，快闪存储器，磁盘或光盘。

处理器 152，与存储器 151 耦合，执行存储器 151 所存储的程序，以用于：

获取本轮的用户语句；从一个或多个候选表格中确定与所述用户语句匹配的一个或多个目标表格；对所述本轮的用户语句进行解析，得到第一查询语句；根据所述第一查询语句，从所述一个或多个目标表格中查询获得目标数据；根据所述目标数据生成本轮的系统回复。

进一步，如图 9 所示，电子设备还可以包括：通信组件 153、电源组件 154、音频组件 155、显示器 156 等其它组件。图 9 中仅示意性给出部分组件，并不意味着电子设备只包括图 9 所示组件。

通信组件 153 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 153 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信组件 153 还包括近场通信(NFC)模块，以促进短程通信。例如，在 NFC 模块可基于射频识别(RFID)技术，红外数据协会(IrDA)技术，超宽带(UWB)技术，蓝牙(BT)技术和其他技术来实现。

电源组件 154，为电子设备的各种组件提供电力。电源组件 154 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 155 被配置为输出和/或输入音频信号。例如，音频组件 155 包括一个麦克风(MIC)，当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦

克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 151 或经由通信组件 153 发送。在一些实施例中，音频组件 155 还包括一个扬声器，用于输出音频信号。

显示器 156 包括屏幕，其屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。

另外，本公开实施例还提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行以实现上述实施例所述的人机对话方法。

本公开实施例还提供了一种计算机程序产品，计算机程序产品包括：计算机程序，计算机程序存储在可读存储介质中，电子设备的至少一个处理器可以从可读存储介质读取计算机程序，至少一个处理器执行计算机程序使得电子设备执行上述任一方法实施例所提供的方案，具体功能和所能实现的技术效果此处不再赘述。

需要说明的是，在本文中，诸如“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

以上所述仅是本公开的具体实施方式，使本领域技术人员能够理解或实现本公开。对这些实施例的多种修改对本领域的技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本公开的精神或范围的情况下，在其它实施例中实现。因此，本公开将不会被限制于本文所述的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

权 利 要 求 书

1、一种人机对话方法，其中，所述方法包括：

获取本轮的用户语句；

从一个或多个候选表格中确定与所述用户语句匹配的一个或多个目标表格；

对所述本轮的用户语句进行解析，得到第一查询语句；

根据所述第一查询语句，从所述一个或多个目标表格中查询获得目标数据；

根据所述目标数据生成本轮的系统回复。

2、根据权利要求 1 所述的方法，其中，根据所述第一查询语句，从所述一个或多个目标表格中查询获得目标数据，包括：

根据历史对话数据中用户历史语句对应的历史查询语句、所述本轮的用户语句对应的会话动作、所述历史对话数据对应的系统状态中的至少一个，对所述第一查询语句进行处理，得到第二查询语句；

根据所述第二查询语句，从所述一个或多个目标表格中查询获得目标数据；

其中，所述会话动作是根据所述历史查询语句和所述第一查询语句得到的；所述历史对话数据对应的系统状态用于表征所述历史对话数据中的系统历史回复是否为反问语句。

3、根据权利要求 1 或 2 所述的方法，其中，根据所述目标数据生成本轮的系统回复，包括：

根据所述目标数据和所述本轮的用户语句，生成所述本轮的系统回复。

4、根据权利要求 1-3 中任一项所述的方法，其中，对所述本轮的用户语句进行解析，得到第一查询语句，包括：

将所述一个或多个目标表格、以及所述本轮的用户语句作为预设模型的输入，通过所述预设模型得到第一查询语句。

5、根据权利要求 4 所述的方法，其中，将所述一个或多个目标表格、以及所述本轮的用户语句作为预设模型的输入，通过所述预设模型得到第一查询语句之前，所述方法还包括：

生成所述一个或多个候选表格中每个候选表格相关的样本查询语句；

将所述样本查询语句转换为第一自然语言语句；

根据所述每个候选表格相关的样本查询语句、以及所述第一自然语言语句，得到训练数据；

根据所述训练数据进行模型训练得到所述预设模型。

6、根据权利要求 5 所述的方法，其中，根据所述训练数据对所述预设模型进行训练，包括：

根据所述训练数据对预训练模型进行训练得到所述预设模型，所述预训练模型是预先根据表格信息训练得到的模型。

7、根据权利要求 5 所述的方法，其中，根据所述每个候选表格相关的样本查询语句、以及所述第一自然语言语句，得到训练数据，包括：

将所述第一自然语言语句中的关键词替换为所述关键词对应的同义词或同义短语，得到所述样本查询语句对应的一个或多个第二自然语言语句；

根据所述每个候选表格相关的样本查询语句、所述第一自然语言语句、以及所述一个或多个第二自然语言语句，得到所述训练数据。

8、一种人机对话装置，其中，所述装置包括：

获取模块，用于获取本轮的用户语句；

确定模块，用于从一个或多个候选表格中确定与所述用户语句匹配的一个或多个目标表格；

解析模块，用于对所述本轮的用户语句进行解析，得到第一查询语句；

查询模块，用于根据所述第一查询语句，从所述一个或多个目标表格中查询获得目标数据；

生成模块，用于根据所述目标数据生成本轮的系统回复。

9、一种电子设备，其中，包括：

存储器；

处理器；以及

计算机程序；

其中，所述计算机程序存储在所述存储器中，并被配置为由所述处理器执行以实现如权利要求 1-7 中任一项所述的方法。

10、一种计算机可读存储介质，其上存储有计算机程序，其中，所述计算机程序被处理器执行时实现如权利要求 1-7 中任一项所述的方法。

11、一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现权利要求 1-7 中任一项所述的方法。